(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
9 January 2003 (09.01.2003)

PCT

(10) International Publication Number
WO 03/003639 A2

(51) International Patent Classification[7]: H04L

(21) International Application Number: PCT/IB02/02404

(22) International Filing Date: 25 June 2002 (25.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/892,611 28 June 2001 (28.06.2001) US

(71) Applicant: NOKIA, INC. [US/US]; 6000 Connection Drive, Irving, TX 75039 (US).

(72) Inventors: CHASKAR, Hemant, M.; 111 Locust Street, Apt. 40-C-1, Woburn, MA 01801 (US). TROSSEN, Dirk; 300 Mystic Valley Parkway, Arlington, MA 02474 (US). KRISHNAMURTHI, Govind; 276 Massachusetts Avenue, #105, Arlington, MA 02474 (US).

(74) Agent: WRIGHT, Bradley, c.; Banner & Witcoff, Ltd., 1001 G Street, N.W., Eleventh Floor, Washington, DC 20001-4597 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/003639 A2

(54) Title: PROTOCOL TO DETERMINE OPTIMAL TARGET ACCESS ROUTERS FOR SEAMLESS IP-LEVEL HANDOVER

(57) Abstract: An apparatus and method is provided for facilitating the seamless handoff of IP connections between access routers in an IP network. The mobile IP network includes two or more access routers each serving a different geographic service area. When a mobile terminal moves from the first service area to the second service area, the mobile terminal transmits to the second access router the IP address of the previous access router. The second access router uses this information to learn capabilities of the first access router (eg., bandwidths supported, security schemes, and the like, for use in future handoff decisions, and exchanges capability information with the first access router. The assumption is made based on the exchanged information that the access routers are geographically proximate. When another mobile terminal transitions from one service area to another, the system selects an optimal target access router based on the previously learned information, including the inferred geographic proximity between access routers.

PROTOCOL TO DETERMINE OPTIMAL TARGET ACCESS ROUTERS FOR SEAMLESS IP-LEVEL HANDOVER

FIELD OF THE INVENTION

[01] The invention relates generally to telecommunications networks. More particularly, the invention concerns a mechanism for enabling seamless mobility in mobile telecommunications networks.

BACKGROUND OF THE INVENTION

[02] Telecommunication networks for mobile devices include cellular communication systems; mobile Internet Protocol (IP) networks; paging systems; and others. Cellular systems generally allow mobile terminals to move geographically by "handing off" localized communication links among transmission towers and associated base stations. Similarly, mobile IP networks allow IP-enabled devices such as wireless Personal Digital Assistants (PDAs) and mobile computers to move about geographically dispersed areas while maintaining a connection to the Internet.

[03] Mobile devices can provide both voice-based connections and IP connections using different base stations and infrastructures. For example, a Web-enabled cell phone might maintain a voice connection using a first transmission channel and maintain a mobile IP connection using a second (and independent) transmission channel, such that handoffs occur independently for the two channels. Alternatively, voice services can be combined with the IP service, such that a single connection is maintained for both services. Voice connections can also be provided over IP in a combined service.

[04] FIG. 1 shows a conventional mobile IP network that covers three service areas SA1, SA2, and SA3. For the sake of simplicity, only IP services are shown, although as explained above, separate transmission networks can be provided for voice services.

[05] As shown in FIG. 1, a mobile terminal MT is within service area SA1 served by base station BS1 (also called an access point or AP). Base station BS1 is connected to an access router AR1, which in turn connects to an Internet service provider ISP1 that provides access to the Internet. Other base stations such as BS3 may also be connected to access router AR1, such that a common IP address is used for mobile terminals even though the terminals may pass through different service areas. In other words, although there may be a hand off of radio frequency channels when the mobile terminal moves between service area SA1 and service area SA3, it may not be necessary to change the IP address used to communicate with the mobile terminal because the Internet connection is still served by the same access router AR1.

[06] A second service area SA2 is served by a separate base station BS2, which is in turn connected to a different access router AR2. Due to the network topology, access routers AR1 and AR2 use different blocks of IP addresses for communicating with mobile terminals roaming within their associated service areas. If mobile terminal MT moves from service area SA1 to service area SA2, some mechanism is needed to hand off the Internet connection from access router AR1 to access router AR2. Similarly, if service areas SA1 and SA2 are separated by a large logical distance (e.g., AR1 and AR2 are connected to different ISPs), some coordination mechanism is needed to permit data transmitted to a terminal previously operating in service area SA1 to be forwarded to service area SA2 if that terminal moves into area SA2.

[07] One conventional scheme for handing off IP connections is depicted in FIG. 2. Service area SA1 is served by access router AR1, which is designated the "home agent" for communicating with a particular mobile terminal MT. While mobile terminal MT moves within service area SA1, access router AR1 communicates with the mobile terminal using an IP address that is assigned to access router AR1. IP packets (e.g., e-mail, Web pages, and the like) are transmitted over the Internet to ISP1, which forwards the traffic to AR1, which in turn knows that a particular IP connection is associated with the mobile terminal in its service area.

[08] If mobile terminal MT moves to a different service area SA2 served by a different access router AR2, packets that were previously transmitted to AR1 will no longer reach the mobile terminal. One conventional solution is to advertise (e.g., broadcast) the existence of access router AR2 in service area SA2, such that when mobile terminal MT moves into service area SA2, it is notified of the existence of access router AR2, and it receives a new IP address for communicating within service area SA2. Mobile terminal MT or access router AR2 then sends a binding update to home agent AR1 (e.g., through a land line LL or over the Internet), so that home agent AR1 knows the IP address that will allow packets to reach the mobile terminal in service area SA2. The home agent treats this address as a "care of" address, and all further packets to the original IP address are forwarded to the new IP address. In essence, two separate IP addresses are used to communicate with the mobile terminal: a home agent address and a care of address that changes at each new point of attachment. This scheme is described in the Internet Engineering Task Force (IETF) Request for Comments (RFC) number 2002 (October 1996).

[09] The above scheme assumes that the target access router (AR2) is known by the originating access router (AR1) prior to the handoff (e.g., mobile terminal MT has accepted the advertisement from AR2 and is assigned an IP address for communicating with it). If there are multiple access routers in the target area each with overlapping service areas, there is no easy way for the mobile terminal to select from among them. For example, suppose that a mobile terminal is receiving high bandwidth video data while moving out of a service area. Two other overlapping service areas served by two access routers controlled by two different service providers may be available to accept the handoff of the mobile terminal's IP connection. One of the two access routers may provide high-speed access to the Internet, while the second one may not. There is no way for the mobile terminal to specify or select from among the two access routers.

[10] Another problem concerns handoff speed. The conventional scenario shown in FIG. 2 may not be able to provide fast handoff speed because of the handshaking required

between the mobile terminal and the new access router AR2. Packets may be lost if handoff of the IP connection is not performed smoothly. Moreover, if an IP connection is used for voice-quality signals or music, latency introduced by the handoff may unacceptably disrupt the connection.

[11] Another difficulty with handing off IP connections in mobile networks arises where heterogeneous networks (using different access technologies) served by potentially different (and incompatible) service providers are concerned. Referring again to FIG. 1, if service area SA1 is served by MCI while service area SA2 is served by AT&T, then the two service providers must agree on a coordination mechanism to accept handoffs of IP services from each other's system. Moreover, as new access routers are added to each service provider's system, the details of each new access router must be communicated throughout the system (e.g., from a central authority) to ensure that all access routers in both systems are aware of the others. This approach can result in a single point of failure, and requires coordination of effort among different service providers.

[12] The problem of providing seamless handovers in IP environments is related to ongoing efforts in the Internet Engineering Task Force (IETF), namely in the Context Transfer, Handoff Candidate Discovery, and Seamless Mobility (SeaMoby) and Mobile IP working groups. Context transfer and fast handover protocols have been developed to exchange session-related information or proactively establish mobile IP connectivity, respectively. Both protocols assume that the target access router is known when requesting the desired functionality (see FIG. 1). Although the discovery of the handoff candidate is included in the SeaMoby working group charter, discovery protocols for physically adjacent access routers have not been studied so far. However, research regarding obtaining physical locations of networking elements has been conducted. Location tracking technologies, such as the Global Positioning System (GPS), provide physical location information of devices attached to the positioning system. Other systems use such information to accurately locate devices. However, since the location is not in relation to any coverage area of an access

technology, the location information is not applicable for candidate selection purposes.

[13] Location systems based on radio frequency technologies use the signal of the wireless access technology to determine the position of the mobile node. In contrast to GPS systems, the obtained location is related to the coverage area of the base station being used for location determination. However, the obtained location is specific for the mobile node and does not give any indication of overlapping coverage areas of access routers. Thus, these systems cannot be used to determine physically adjacent networking elements. Moreover, the location determination is usually very specific for the access technology used, and is therefore not suited for multiple access technology scenarios. Besides the lack of accuracy of the obtained location, there is no indication of overlapping coverage areas needed for physical adjacency determination.

[14] What is needed is a system and method for addressing some or all of the aforementioned problems.

SUMMARY OF THE INVENTION

[15] The invention provides a system and method to facilitate seamless handoffs in mobile networks, such as mobile IP networks. A first aspect of the invention enables an access router to dynamically learn about other access routers that are geographically adjacent by receiving information from mobile terminals that move into the service area of the access router. A second aspect of the invention allows access routers to share capability information without requiring a centralized scheme (e.g., using a peer-to-peer approach). A third aspect of the invention allows a target access router to be selected and a handoff arranged on the basis of capability information associated with one or more target access routers and on the basis of the direction of movement of the mobile node. Other features and advantages of the invention will become apparent through the following detailed description, the figures, and the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[16] FIG. 1 shows a conventional mobile IP network covering three service areas SA1, SA2, and SA3.

[17] FIG. 2 shows a conventional scheme for handing off IP connections, in which a mobile terminal registers with a home agent AR1 but also communicates using a second IP address through a "care of" agent AR2.

[18] FIG. 3 shows a system according to the invention including a plurality of access routers AR1 and AR2, each of which includes a capability map (304 and 308) describing capabilities of geographically proximate access routers.

[19] FIG. 4 shows steps in a method to learn about a physical neighborhood and for acquiring capability information from physically adjacent access routers.

[20] FIG. 5 shows steps in a method for selecting a target access router on the basis of previously stored capability information.

[21] FIG. 6 shows a mobile terminal MT moving from a service area associated with a first access router AR1 to an area serviced by three different access routers AR2, AR3, and AR4.

[22] FIG. 7 shows a mobile terminal 701 equipped with processing functions and memory to carry out various aspects of the invention.

DETAILED DESCRIPTION OF THE INVENTION

[23] FIG. 3 shows a system employing various principles of the invention. As shown in FIG. 3, a first access router AR1 serves a first service area (not shown) in which a mobile terminal MT may be located. Although not explicitly shown in FIG. 3, it is assumed that each access router transmits and receives data packets through one or more base stations that cover corresponding geographic areas. It is also assumed that

each access router provides Internet-compatible connections (e.g., IP protocol compatibility) such that data packets received at each router can be forwarded to one or more mobile terminals within the corresponding service area. Each access router includes an IP address used for communicating directly with the access router and a block of IP addresses that can be allocated and used by the access router for communicating with mobile terminals served by the access router. For purposes of illustration, AR1 is shown as having an IP address of 10.1.0.0, and AR2 is shown as having an IP address of 10.2.0.0.

[24] According to one aspect of the invention, each access router creates and maintains a local capability map (elements 304 and 308 in FIG. 3) that stores information concerning other access routers that are geographically adjacent. According to one aspect of the invention, as a mobile terminal MT moves into the area serviced by an access router, the mobile terminal transmits the IP address of the access router for the service area from which the mobile terminal is leaving. In other words, each mobile terminal passes to the next access router information concerning the previously used access router (the previous router's identity, i.e., its IP address). An inference can be drawn that, by virtue of moving out of one router's service area and into another router's service area, the two routers are geographically adjacent. Once each access router knows about the other one, they can exchange capability information that can be used to select a target access router for future handoffs. The capability information, along with the physical neighborhood AR map, can also be constructed through manual configuration.

[25] As shown in FIG. 3, access router AR1 includes a learning function 301, a selector function 302, and an exchange function 303. Similarly, access router AR2 contains such functions (elements 305, 306, and 307) in addition to the capabilities map 308. Other access routers AR3 and AR4 are shown without internal details. In general, each learning function 301 and 307 receives information from mobile terminals that move into the service area associated with an access router (e.g., the IP address of the previously used access router).

[26] Exchange functions 303 and 305 exchange capability information between two access routers in response to the learning function. For example, when mobile terminal MT is about to move out of the service area supported by AR1 and into the service area of AR2, the mobile terminal transmits to AR2 the IP address (in this case, 10.1.0.0) of the originating access router AR1. In response, learning function 307 stores the IP address of AR1 into capability map 308, and causes exchange function 305 to transmit a request (over the Internet, or through other means) to AR1 to exchange capability information. Thereafter, exchange functions 303 and 305 of the respective access routers exchange capability information (described in more detail below) concerning each respective router's capabilities. For example, if AR1 can support link bandwidths of 28 KBPS and AR2 can support link bandwidths of 56 KBPS, this information is stored in each access router's respective capability map. In this manner, each access router learns about capabilities of neighboring routers.

[27] Selector functions 302 and 306 select target access routers for mobile terminals based on capability information stored in capability maps 304 and 308 respectively. For example, if mobile terminal MT is about to move from a service area served by AR1 into a service area served by multiple target access routers (including, for example, AR2 and AR4), selector function 302 in AR1 consults capability map 304 to determine which access router best suits the capabilities needed by mobile terminal MT. A movement detection scheme is used to inform AR1 which ARs are reachable by the mobile terminal upon movement of the mobile terminal. As explained in more detail below, selection of target routers can be done based on policies stored in each router.

[28] Capabilities associated with each access router may include static capabilities (e.g., bandwidths supported by the router; security protocols; service providers; etc.) and dynamic capabilities (e.g., current loading level or network delays). Examples shown in FIG. 3 include bandwidths supported; security schemes; ISP connected to the router; IP address of the router; quality of service parameters; and dynamic loading conditions.

[29] Any or all of the functions depicted in FIG. 3 can be implemented using computer software executing on a general-purpose or special-purpose digital computer. The capabilities information can be stored in a computer memory, relational database, or other data structure. Conventional access routers can be modified to incorporate the functions illustrated in FIG. 3.

[30] Suppose that the user of mobile terminal MT is watching a movie over an IP connection requiring a connection bandwidth of 256 kilobytes per second (KBPS). Suppose further that terminal MT is about to move from an access router that presently supports such a bandwidth to an area served by two access routers, AR2 and AR4. The movement detection scheme allows AR1 to know that AR2 and AR4 can cover MT after it moves out of service area SA1. Selector function 302 in access router AR1 consults capability map 304 and determines that of the two access routers in the area, only AR4 supports such a bandwidth. Thereafter, AR1 arranges a handoff between mobile terminal MT and access router AR4. Arranging a handoff may include procedures of context transfer (see, e.g., R. Koodli and C. Perkins, "A Context Transfer Framework for Seamless Mobility," Work in Progress, Internet Draft, February 2001), or fast handover (see, e.g., G. Tsirtsis et al., "Fast Handovers for Mobile IPv6," Work in Progress, Internet Draft, April 2001).

[31] A handoff can be arranged in various ways, including instructing the MT to contact AR4; sending a message to AR4 to arrange the handoff; or by other means. It will be appreciated that the selection function can be performed in another router, processor, or mobile terminal.

[32] Suppose that the user of mobile terminal MT requires a high-security connection that supports 128-bit encryption. When terminal MT moves from one service area to another, it informs the old access router of the list of reachable access routers. The old access router selects a target router based on the MT's requirements and the stored capability information of the appropriate neighbor routers. The selection process is explained in more detail below. Other selection schemes can of course be used.